

RFC 4893 : BGP Support for Four-octet AS Number Space

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 22 Mai 2007

Date de publication du RFC : Mai 2007

<http://www.bortzmeyer.org/4893.html>

Un des nombres trop petits de l'Internet était la taille des numéros de système autonome. Elle passe, avec ce RFC, de deux à quatre octets.

Il est assez courant dans l'Internet que des nombres prévus très larges au début s'avèrent ridiculement petits avec la croissance du réseau. C'est bien sûr le cas des adresses IPv4, dont les 32 bits sont bien trop peu pour le nombre de machines connectées aujourd'hui, mais c'est aussi vrai pour les numéros de système autonome (AS pour "*autonomous system*"). Chacun de ces numéros identifie un système autonome de routage, au sein duquel une politique de routage cohérente peut s'appliquer. En gros, chaque opérateur a un numéro de système autonome, les plus gros en ayant plusieurs (surtout en cas de fusion ou d'acquisition).

Ces numéros sont notamment utilisés par le protocole de routage BGP (normalisé dans le RFC 4271¹), pour indiquer le chemin à suivre pour joindre un réseau. La petite fonction shell (en ligne sur <http://www.bortzmeyer.org/files/bgproute>) permet d'interroger les serveurs de Route Views <<http://routeviews.org>> et affiche les numéros de système autonomes traversés :

```
% bgproute 80.67.162.1
AS path: 3257 3356 20766 20766
Route: 80.67.160.0/19
```

1. Pour voir le RFC de numéro NNN, <http://www.ietf.org/rfc/rfcNNN.txt>, par exemple <http://www.ietf.org/rfc/rfc4271.txt>

On voit que le serveur de Route Views a reçu l'annonce de la part du système autonome **3257** qui l'avait lui même reçu du système autonome **3356**.

Ces numéros étaient stockés sur seulement 16 bits, ce qui ne permettait que 65 535 systèmes en tout, bien trop peu pour l'Internet d'aujourd'hui, qui va des villes chinoises aux hébergeurs brésiliens. Si certains conservateurs, méprisants et élitistes, ont regretté le fait que « n'importe qui, avec une armoire et deux PC avec Quagga » veuille faire du BGP, le fait est que l'Internet touche bien plus de monde et que la population des opérateurs a augmenté. D'où notre RFC, qui fait passer la taille des numéros d'AS à 32 bits, soit quatre milliards d'opérateurs possibles.

Ces nouveaux AS s'écriront `<http://www.bortzmeyer.org/representation-texte.html>` en notation `ASPLAIN`, en écrivant directement le nombre, par exemple `112617` (RFC 5396).

Le changement lui-même est assez trivial mais, comme souvent sur Internet, le gros du travail (et qui prend la plus grande partie de notre RFC) était la gestion de la transition. Notre RFC explique avec beaucoup de soin comment un routeur BGP récent va pouvoir parler à un routeur de l'ancienne génération et comment les chemins d'AS 4-octets pourront être transmis même à travers des « vieux » routeurs, utilisant un mécanisme de tunnel (l'article de Geoff Huston `<http://www.potaroo.net/ispcol/2007-01/asn32.html>` l'explique très bien).

Pour cette transition, le nouveau BGP utilise un numéro d'AS spécial, le **23456**, qui sert à représenter tous les AS 4-octets pour les anciens routeurs. Si vous voyez apparaitre ce système autonome, par exemple en tapant un `show ip bgp` sur un Cisco, c'est que votre logiciel est trop vieux.

La publication de notre RFC ne fait qu'entériner un changement déjà bien entamé ; cela fait plusieurs mois que les RIR allouaient des numéros sur quatre octets (voir par exemple la politique du RIPE-NCC `<http://www.ripe.net/ripe/docs/asn-assignment.html#19>`).

Ainsi, le système autonome `196613` est utilisé depuis plusieurs mois pour des tests. Le routeur de Route Views étant de la vieille génération, le préfixe `145.125.0.0/20`, annoncé par l'AS `196613`, est vu comme émanant du célèbre `23456` :

```
% bgproute 145.125.0.1
AS path: 14608 19029 3356 3549 1103 1125 23456
Route: 145.125.0.0/20
```

Un petit regret personnel : l'encodage des paquets BGP étant insuffisamment robuste, rien n'indique à un programme qui analyse une session BGP (par exemple Wireshark) que la session utilise des numéros à quatre octets et on ne peut donc pas analyser une session si on n'a pas vu l'échange initial, où s'est décidé de coder les numéros d'AS sur deux ou quatre octets. Comme les sessions BGP durent souvent des semaines, cela peut être un problème gênant.